The Continuity MODIS-VIIRS Cloud Mask (MVCM) User Guide

Based on NASA MODIS Cloud Mask (MOD35, MYD35)
Reprocessed Data – Version 1
Product User Guide – Version 1.0

Richard Frey, Steve Ackerman, Robert Holz, Steve Dutcher
Space Science and Engineering Center
University of Wisconsin-Madison
February 2019

Table of Contents

Introduction

The Continuity MODIS-VIIRS cloud mask (MVCM) is designed to facilitate continuity in cloud detection between the MODIS (Moderate Resolution Imaging Spectroradiometer) on the Aqua and Terra platforms and the series of VIIRS (Visible Infrared Imaging Radiometer Suite) instruments, beginning with the Suomi NPP spacecraft. Global Terra and Aqua MODIS cloud products, including cloud fraction, are available from early 2000 until present and are expected to continue through 2022. The VIIRS data begins in early 2012 and will potentially be available through 2040 on various satellite platforms. Together, these instruments would help generate an unparalleled 40-year record of satellite-derived ocean, land, and atmosphere measurements.

MODIS instruments measure upwelling radiation from the earth-atmosphere system in 36 spectral bands (Ackerman, et al., 2010), while the VIIRS has a subset of these in 20 bands, plus a day/night visible channel (Xiong, et. al., 2014). To facilitate continuity, the MVCM utilizes only those channels that are common to both instruments (see Table 1). Pixel by pixel, clear- vs. cloudy-sky discrimination is accomplished using the same "fuzzy logic" methodology as the MODIS cloud mask (MOD35, Ackerman, et al., 1998, 2010). Details of the algorithm and the history of its development may be found in Ackerman, et al., 2010. Nominal spatial resolution of the MVCM L2 product is 1-km for MODIS and 750-m for VIIRS. The MVCM is capable of processing both MODIS and VIIRS inputs; in this document we refer to VIIRS output from the algorithm as "MVCM VIIRS" and MODIS output as "MVCM Aqua MODIS". Terra MODIS inputs have not yet been produced by the MVCM.

Table 1. MODIS and VIIRS spectral bands used in the MVCM

Spectral Bands Used in the MODIS-VIIRS Cloud Mask (MVCM)			
MODIS Wavelengths (µm)	**MODIS Band**	**VIIRS Band**	**Primary Use**
0.412	8	M1	daytime desert cloud detection
0.443	9	M2	sun glint clear sky detection
0.555	4	M4	snow/ice detection
0.645	1	M5	land surface cloud detection
0.859	2	M7	water surface cloud detection
1.24	5	M8	turbid water clear sky detection
1.375	26	M9	transmissive cirrus cloud detection
1.64	6	M10	snow/ice detection, water surface cloud detection
2.13	7	M11	snow/ice detection, water surface cloud detection
3.75	20	M12	land and water surface cloud detection (VIIRS)
3.96	21	not used	land and water surface cloud detection (MODIS)
8.55	29	M14	water surface ice cloud detection
11.03	31	M15	night land and water surface cloud detection
12.02	32	M16	transmissive cirrus cloud detection

Transition from MOD35 to MVCM

Clouds are generally brighter and colder than their underlying surfaces. Therefore, during daylight hours a majority are discernable by using visible and near-infrared (VNIR) reflectances along with longwave infrared (LWIR) measurements. (Here we use LWIR to mean the atmospheric window region from 8-12 µm.) At night, LWIR brightness temperatures are sufficient to detect most middle and high altitude clouds. However, it is in areas with reduced VNIR contrast during the day (e.g., ocean sun glint) and reduced LWIR contrast at night (e.g., oceanic low altitude clouds, polar night) that infrared measurements in atmospheric gas absorbing spectral regions become important. Fewer of these bands in VIIRS instrument compared to MODIS accounts for most of the differences between the MVCM and MOD35 algorithms.

The most important atmospheric absorption bands for MODIS cloud detection are the water vapor absorption channels at 6.7 and 7.3 µm. Additionally, the CO_2 absorption bands at 13.3 and 13.9 µm play a significant role in detection of clear skies in polar night conditions. (These "clear-sky restoral" tests are performed to find unambiguously clear pixels for certain scene types, and are not strictly part of the "fuzzy logic" algorithm.) Also, MODIS bands 17 and 18 in the 0.9 µm water vapor absorption band are used in clear-sky restoral tests for sun glint conditions. Table 2 shows MODIS spectral bands and cloud tests used in MOD35 that are not found in the MVCM.

Table 2. Spectral bands and uses in MOD35 that are not part of the MVCM

MODIS Spectral Cloud and Clear-sky Tests Not Found in the MVCM		
Wavelengths (µm)	**MODIS Band**	**Use in MOD35**
0.905	17	Clear-sky detection in sun glint conditions (0.905 / 0.936 µm)
0.936	18	Clear-sky detection in sun glint conditions (0.905 / 0.936 µm)
6.7	27	Global high cloud BT* threshold test; clear-sky detection in polar night conditions (6.7–11 µm BTD**)
7.3	28	Nighttime middle cloud detection over land, polar night cloud detection, polar night clear-sky detection (7.3–11 µm BTD); nighttime ocean low cloud detection (8.6–7.3 µm BTD)
13.3	33	Clear-sky detection in polar night conditions (13.3–11 µm BTD)
13.9	35	Mid-latitude (60S–60N) high cloud BT threshold test; clear-sky detection in polar night conditions (13.9–11 µm BTD)

*BT = Brightness temperature; **BTD = Brightness temperature difference

New Spectral Tests in the MVCM

In an attempt to make up for this loss of information due to the lack of the absorption bands in the VIIRS, and to take advantage of new algorithm development, there are several new features in the MVCM that are currently not found in MOD35 algorithm.

1.6/2.1 µm ocean day threshold test

A new threshold test employs reflectances from the 1.6 µm (VIIRS and Terra MODIS) and 2.1 µm (Aqua MODIS) bands to better detect water phase clouds over daytime water surfaces. The very dark ocean background in these spectral bands is especially helpful for thin clouds, partially cloud-filled pixels, and cloud edges.

Turbid water test

A new turbid water clear-sky test has been implemented for shallow waters, and follows the method of Chen and Zhang (2015). Bottom and suspended sediments in near-shore waters can have strong reflectance signals in VNIR bands, and result in false cloud determinations. Reflectance standard deviations at 2.1 µm are calculated over the 3x3 pixel regions centered on the pixels of interest. Pixels in regions with smaller standard deviations than the threshold value are labeled as *clear*.

Test for snow cover over vegetated regions

An addition to the normalized difference snow index (NDSI) test has been added, following Klein et. al., (1998). For vegetated scenes where the NDSI is lower than expected for snow cover, a normalized difference vegetation index (NDVI) value is compared to a threshold that is itself a function of the NDSI. If the NDVI is less than the calculated threshold, snow cover is assumed.

New Test Thresholds in the MVCM

A general strategy for transitioning cloud tests from MOD35 to the MVCM was to "tighten up" thresholds, i.e., tune the tests such that they detect as many clouds as possible without greatly increasing false positives. This was done in order to make up for lesser amounts of information in the VIIRS spectral measurements. For a complete discussion of thresholds and thresholding methods used in the MVCM and MOD35, see the MOD35 Algorithm Theoretical Basis Document (ATBD, Ackerman et al., 2010, online at https://modis-atmosphere.gsfc.nasa.gov/documentation/atbds-plans-guides). The following changes were made to MVCM cloud test thresholds as compared to MOD35.

Daytime land 1.38 µm cirrus test

These thresholds were lowered from {0.04, 0.035, 0.030} to {0.0375, 0.0250, 0.0125} for low, middle, and high confidence of clear sky, respectively.

Daytime water 0.86, 1.6/2.1, and 1.38 µm thresholds

In the MVCM, these thresholds are functions of solar zenith angle (SZA) in the following form:

$$thr = coeff[0] + coeff[1]*sza + coeff[2]*sza^2 + coeff[3]*sza^3 + \ldots ,$$

where "***thr***" is the high confidence clear-sky threshold. The middle and low confidence thresholds are calculated as offsets from "***thr***". Separate coefficients exist for MODIS and VIIRS, and both were taken

from collocated imager reflectances and CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) lidar data, where the CALIOP cloud product served as "truth" for clear vs. cloudy skies. The thresholds are further defined according to the viewing zenith angle (VZA) in the following form:

$$thr_{vza} = thr * (1.0 / \cos(vza)^{p}),$$

where ***p*** = 0.75 for VIIRS and 0.50 for MODIS (0.86 and 1.38 µm). The value of ***p*** for VIIRS 1.6 µm is 0.25; no VZA adjustment is used for MODIS Aqua (2.1 µm). Additional upward adjustments are made for the 1.38 µm thresholds beyond 45 degrees SZA (maximum of 0.02 at 90 degrees SZA). This test will need monitoring because of the future possibility of more open water at polar latitudes. Very low values of atmospheric moisture in these regions make adjustments mandatory.

Inputs to the MVCM

Level-1b Calibrated Radiance Data

The MVCM obtains standard calibrated and geolocated radiance data from both Aqua MODIS and VIIRS data streams (from NASA) but several additional steps are required. For MODIS, bands 1 (0.645 µm) and 2 (0.859 µm) at 250 m spatial resolution are "reaggregated", accounting for focal plane misalignment between the 250 m and 500 m spectral bands. In addition, a destriping algorithm based on Weinreb et al., (1989) is applied to bands 20 and 22–36. It accounts for both detector-to-detector, and mirror-side striping, resulting in significant reduction in striping noise. For SNPP VIIRS, a bias correction is made for bands M5, M7, M8, M10, and M11. The biases were computed via comparisons to Aqua MODIS reflectances, and are applied through scaling factors when reading the L1b files.

Ancillary Data

Several ancillary data sets serve as inputs to the MVCM process, which are listed and briefly described below.

- Global Data Assimilation System (GDAS) files contain gridded model output products including surface temperature and total precipitable water used by the MVCM. They include four files per day at UTC times of 00, 06, 12, and 18Z.
- Near-Real-Time SSM/I-SSMIS (Special Sensor Microwave/Imager-Special Sensor Microwave Imager/Sounder) EASE (Equal-Area Scalable Earth)-Grid Daily Global Ice Concentration and Snow Extent (NISE) files contain daily global gridded snow and ice extent.
- Optimum Interpolated Sea Surface Temperature (OISST) files contain weekly ocean surface temperatures.
- Olson ecosystem static files contain global high spatial resolution ecosystem types (indices 0–99).
- Normalized Difference Vegetation Index (NDVI) maps from Moody et al., (2005) provide gap-filled, global high spatial resolution background NDVIs at 16-day intervals throughout the calendar year.
- MODIS-derived land-water mask

Output Product Files

Contents

Output MVCM files are in netCDF4 format. In addition to cloud mask results, they contain geolocation information, scan line start-times, and various attribute data. Figure 1 shows a header dump of an output file from July 6, 2014 at 17:00 UTC. Array data is divided into three groups: scan line attributes, geolocation data, and geophysical data. The cloud mask results reside in the geophysical data group. The 'Cloud_Mask' (CM) and 'Quality_Assurance' (QA) arrays are the same as in MOD35 except that the bits indicating 250-m results are empty (=0). Bits that indicate whether or not a particular spectral test was performed are included in QA, and are found in the same bit locations as in the CM (see Table 3). Other QA information is detailed in the "MODIS Atmosphere QA Plan for Collection 6" (see https://modis-atmosphere.gsfc.nasa.gov/sites/default/files/ModAtmo/QA_Plan_C6_Master_2015_05_05_0.pdf).
Except for the change from HDF4 to netCDF4 format, the above arrays may be read with the same or very similar computer code as was used for MOD35. For detailed information about accessing and interpreting these bit flags and some example codes, see Appendix A of the MOD35 ATBD. Bit locations in the CM array have not changed, but as indicated above, some tests found in the MOD35 algorithm are not included in the MVCM, and some others have been added. Table 3 shows the various bit locations for data in the CM array. Note that there are no 250-m (MODIS) or I-band (VIIRS) results reported in the file. MVCM VIIRS data will cover six minutes of time in the along-track direction while MVCM Aqua MODIS will cover five minutes in length as in MOD35. VIIRS data spans more across-scan distance than MODIS ('number_of_pixels' = 3200 instead of 1354).

Two cloud mask output arrays exist in the MVCM files that were not part of MOD35. They include the 'Clear_Sky_Confidence' and 'Integer_Cloud_Mask' data sets. The 'Clear_Sky_Confidence' is the final numeric value of the confidence of clear sky, or Q value (Ackerman et al., 2010), computed by the MVCM algorithm. This Q value is converted into one of four cloud-mask categories (confident clear, probably clear, probably cloudy, confident cloudy) reported in bits 1 and 2 (0-based) of the cloud mask (see Table 3). The other array, 'Integer_Cloud_Mask', is the value of bits 1 and 2 converted to an integer value. This serves users who would like to use the cloud mask categories without unpacking the remaining bits in the CM array. The integer values are 0–3, corresponding to confident cloudy, probably cloudy, probably clear, and confident clear, respectively. A value of -1 indicates no result (fill value).

Figure 1. Header dump of sample output CLDMSK_L2_VIIRS_SNPP netCDF4 file

```
netcdf CLDMSK_L2_VIIRS_SNPP.A2019038.0142.001.2019060173405 {
dimensions:
          number_of_lines = 3232 ;
          number_of_pixels = 3200 ;
          byte_segment = 6 ;
          QA_dimension = 10 ;
          number_of_scans = 202 ;

// global attributes:
                    :format_version = 1 ;
                    :platform = "Suomi-NPP" ;
                    :processing_level = "L2" ;
                    :processing_version = "v1.0" ;
                    :cdm_data_type = "swath" ;
                    :institution = "NASA VIIRS Atmosphere SIPS" ;
                    :keywords_vocabulary = "NASA Global Change Master Directory (GCMD)Science
Keywords" ;
                    :license = "http://science.nasa.gov/earth-science/earth-science-data/data-
information-policy/" ;
                    :stdname_vocabulary = "NetCDF Climate and Forecast (CF) Metadata Convention" ;
                    :naming_authority = "gov.nasa.gsfc.sci.atmos" ;
                    :NCO = "\"4.5.5\"" ;
                    :title = "SNPP VIIRS Cloud Mask and Spectral Test Results
(CLDMSK_L2_SNPP_VIIRS)" ;
                    :long_name = "VIIRS/SNPP Cloud Mask and Spectral Test Results 6-Min L2 Swath
750m" ;
                    :Conventions = "CF-1.6, ACDD-1.3" ;
                    :instrument = "VIIRS" ;
                    :creator_name = "NASA VIIRS Atmosphere SIPS" ;
                    :creator_email = "sips.support@ssec.wisc.edu" ;
                    :creator_url = "https://sips.ssec.wisc.edu" ;
                    :project = "NASA VIIRS Atmosphere SIPS" ;
                    :publisher_name = "LAADS" ;
                    :publisher_email = "modis-ops@lists.nasa.gov" ;
                    :publisher_url = "https://ladsweb.modaps.eosdis.nasa.gov/" ;
                    :history = "" ;
                    :source = "iff-cloud 1.1, mvcm 20190117-1" ;
                    :date_created = "2019-03-01T17:32:45Z" ;
                    :product_name = "CLDMSK_L2_VIIRS_SNPP.A2019038.0142.001.2019060173405.nc" ;
                    :LocalGranuleID = "CLDMSK_L2_VIIRS_SNPP.A2019038.0142.001.2019060173405.nc" ;
                    :ShortName = "CLDMSK_L2_VIIRS_SNPP" ;
                    :product_version = "1.0" ;
                    :AlgorithmType = "OPS" ;
                    :identifier_product_doi = "10.5067/VIIRS/CLDMSK_L2_VIIRS_SNPP.001" ;
                    :identifier_product_doi_authority = "http://dx.doi.org" ;
                    :input_files =
"VNP03MOD.A2019038.0142.001.2019038062600.uwssec.nc,VNP02MOD.A2019038.0142.001.2019038062600.uwss
ec.bowtie_restored_scaled.nc" ;
                    :ancillary_files =
"NISE_SSMISF18_20190206.HDFEOS,oisst.20190130,gdas1.PGrbF00.190207.00z,gdas1.PGrbF00.190207.06z";
                    :l1_version = "2.0.2" ;
                    :l1_lut_version = "2.0.0.28" ;
                    :l1_lut_created = "2019-01-04" ;
                    :DataCenterId = "UWI-MAD/SSEC/ASIPS" ;
                    :creator_institution = "Space Science & Engineering Center, University of
Wisconsin - Madison" ;
                    :publisher_institution = "NASA Level-1 and Atmosphere Archive & Distribution
System" ;
                    :GRingPointSequenceNo = 1LL, 2LL, 3LL, 4LL ;
                    :GRingPointLatitude = -9.66756629943848, -5.41480398178101, 15.2493619918823,
10.9436693191528 ;
                    :GRingPointLongitude = 6.83322048187256, -20.6769676208496, -16.2769317626953,
11.611307144165 ;
                    :geospatial_lat_units = "degrees_north" ;
                    :geospatial_lon_units = "degrees_east" ;
                    :geospatial_lat_min = -9.66756629943848 ;
                    :geospatial_lat_max = 15.2493619918823 ;
                    :geospatial_lon_min = -20.6769676208496 ;
                    :geospatial_lon_max = 11.611307144165 ;
```

```
                    :NorthBoundingCoordinate = 15.2493619918823 ;
                    :SouthBoundingCoordinate = -9.66756629943848 ;
                    :EastBoundingCoordinate = 11.611307144165 ;
                    :WestBoundingCoordinate = -20.6769676208496 ;
                    :time_coverage_start = "2019-02-07T01:42:00.000Z" ;
                    :time_coverage_end = "2019-02-07T01:48:00.000Z" ;
                    :startDirection = "Descending" ;
                    :endDirection = "Descending" ;
                    :OrbitNumber = 37720LL ;
                    :DayNightFlag = "Night" ;
                    :xmlmetadata = "<?xml version=\"1.0\"?>\n<!DOCTYPE GranuleMetaDataFile SYSTEM
\"http://ecsinfo.gsfc.nasa.gov/ECSInfo/ecsmetadata/dtds/DPL/ECS/ScienceGranuleMetadata.dtd\">\n<G
ranuleMetaDataFile>\n  <DTDVersion>1.0</DTDVersion>\n  <DataCenterId>UWI-
MAD/SSEC/ASIPS</DataCenterId>\n  <GranuleURMetaData>\n    <CollectionMetaData>\n
<ShortName>CLDMSK_L2_VIIRS_SNPP</ShortName>\n      <VersionID>1</VersionID>\n
</CollectionMetaData>\n    <ECSDataGranule>\n      <ReprocessingPlanned>no further reprocessing
anticipated</ReprocessingPlanned>\n
<LocalGranuleID>CLDMSK_L2_VIIRS_SNPP.A2019038.0142.001.2019060173405.nc</LocalGranuleID>\n \n
<DayNightFlag>Night</DayNightFlag>\n      <ProductionDateTime>2019-03-01
17:34:05.247459</ProductionDateTime>\n      <LocalVersionID>1</LocalVersionID>\n
</ECSDataGranule>\n    <PGEVersionClass>\n      <PGEVersion>20190117-1</PGEVersion>\n
</PGEVersionClass>\n    <RangeDateTime>\n
<RangeEndingTime>01:48:00.000000</RangeEndingTime>\n      <RangeEndingDate>2019-02-
07</RangeEndingDate>\n      <RangeBeginningTime>01:42:00.000000</RangeBeginningTime>\n
<RangeBeginningDate>2019-02-07</RangeBeginningDate>\n    </RangeDateTime>\n
<SpatialDomainContainer>\n      <HorizontalSpatialDomainContainer>\n        <BoundingRectangle>\n
<WestBoundingCoordinate>-20.6769676208</WestBoundingCoordinate>\n
<NorthBoundingCoordinate>15.2493619919</NorthBoundingCoordinate>\n
<EastBoundingCoordinate>11.6113071442</EastBoundingCoordinate>\n
<SouthBoundingCoordinate>-9.66756629944</SouthBoundingCoordinate>\n        </BoundingRectangle>\n
</HorizontalSpatialDomainContainer>\n    </SpatialDomainContainer>\n
<OrbitCalculatedSpatialDomain>\n      <OrbitCalculatedSpatialDomainContainer>\n
<OrbitNumber>37720</OrbitNumber>\n      </OrbitCalculatedSpatialDomainContainer>\n
</OrbitCalculatedSpatialDomain>\n    <Platform>\n      <PlatformShortName>Suomi
NPP</PlatformShortName>\n      <Instrument>\n
<InstrumentShortName>VIIRS</InstrumentShortName>\n        <Sensor>\n
<SensorShortName>VIIRS</SensorShortName>\n        </Sensor>\n      </Instrument>\n
</Platform>\n    <InputGranule>\n
<InputPointer>VNP03MOD.A2019038.0142.001.2019038062600.uwssec.nc</InputPointer>\n
<InputPointer>VNP02MOD.A2019038.0142.001.2019038062600.uwssec.bowtie_restored_scaled.nc</InputPoi
nter>\n    </InputGranule>\n    <AncillaryInputGranules>\n      <AncillaryInputGranule>\n
<AncillaryInputType>NISE</AncillaryInputType>\n
<AncillaryInputPointer>NISE_SSMISF18_20190206.HDFEOS</AncillaryInputPointer>\n
</AncillaryInputGranule>\n      <AncillaryInputGranule>\n
<AncillaryInputType>REYNSST</AncillaryInputType>\n
<AncillaryInputPointer>oisst.20190130</AncillaryInputPointer>\n      </AncillaryInputGranule>\n
<AncillaryInputGranule>\n        <AncillaryInputType>GDAS_0ZF</AncillaryInputType>\n
<AncillaryInputPointer>gdas1.PGrbF00.190207.00z</AncillaryInputPointer>\n
</AncillaryInputGranule>\n      <AncillaryInputGranule>\n
<AncillaryInputType>GDAS_0ZF</AncillaryInputType>\n
<AncillaryInputPointer>gdas1.PGrbF00.190207.06z</AncillaryInputPointer>\n
</AncillaryInputGranule>\n    </AncillaryInputGranules>\n
</GranuleURMetaData>\n</GranuleMetaDataFile>" ;

group: geolocation_data {
  variables:
          float latitude(number_of_lines, number_of_pixels) ;
                    latitude:long_name = "Latitudes of pixel locations" ;
                    latitude:units = "degrees_north" ;
                    latitude:_FillValue = -999.9f ;
                    latitude:valid_min = -90.f ;
                    latitude:valid_max = 90.f ;
          float longitude(number_of_lines, number_of_pixels) ;
                    longitude:long_name = "Longitudes of pixel locations" ;
                    longitude:units = "degrees_east" ;
                    longitude:_FillValue = -999.9f ;
                    longitude:valid_min = -180.f ;
                    longitude:valid_max = 180.f ;
          short sensor_azimuth(number_of_lines, number_of_pixels) ;
                    sensor_azimuth:long_name = "Sensor azimuth angle at pixel locations" ;
                    sensor_azimuth:units = "degrees" ;
```

```
                    sensor_azimuth:_FillValue = -32768s ;
                    sensor_azimuth:valid_min = -18000s ;
                    sensor_azimuth:valid_max = 18000s ;
                    sensor_azimuth:scale_factor = 0.01f ;
                    sensor_azimuth:add_offset = 0.f ;
            short sensor_zenith(number_of_lines, number_of_pixels) ;
                    sensor_zenith:long_name = "Sensor zenith angle at pixel locations" ;
                    sensor_zenith:units = "degrees" ;
                    sensor_zenith:_FillValue = -32768s ;
                    sensor_zenith:valid_min = 0s ;
                    sensor_zenith:valid_max = 18000s ;
                    sensor_zenith:scale_factor = 0.01f ;
                    sensor_zenith:add_offset = 0.f ;
            short solar_azimuth(number_of_lines, number_of_pixels) ;
                    solar_azimuth:long_name = "Solar azimuth angle at pixel locations" ;
                    solar_azimuth:units = "degrees" ;
                    solar_azimuth:_FillValue = -32768s ;
                    solar_azimuth:valid_min = -18000s ;
                    solar_azimuth:valid_max = 18000s ;
                    solar_azimuth:scale_factor = 0.01f ;
                    solar_azimuth:add_offset = 0.f ;
            short solar_zenith(number_of_lines, number_of_pixels) ;
                    solar_zenith:long_name = "Solar zenith angle at pixel locations" ;
                    solar_zenith:units = "degrees" ;
                    solar_zenith:_FillValue = -32768s ;
                    solar_zenith:valid_min = 0s ;
                    solar_zenith:valid_max = 18000s ;
                    solar_zenith:scale_factor = 0.01f ;
                    solar_zenith:add_offset = 0.f ;
    } // group geolocation_data

group: geophysical_data {
    variables:
            float Clear_Sky_Confidence(number_of_lines, number_of_pixels) ;
                    Clear_Sky_Confidence:long_name = "VIIRS Clear Sky Confidence" ;
                    Clear_Sky_Confidence:units = "none" ;
                    Clear_Sky_Confidence:_FillValue = -999.9f ;
                    Clear_Sky_Confidence:valid_min = 0.f ;
                    Clear_Sky_Confidence:valid_max = 1.f ;
            ubyte Cloud_Mask(byte_segment, number_of_lines, number_of_pixels) ;
                    Cloud_Mask:long_name = "VIIRS Cloud Mask and Spectral Test Results" ;
                    Cloud_Mask:units = "none" ;
                    Cloud_Mask:_FillValue = 0UB ;
                    Cloud_Mask:valid_min = 1 ;
                    Cloud_Mask:valid_max = 255 ;
            byte Integer_Cloud_Mask(number_of_lines, number_of_pixels) ;
                    Integer_Cloud_Mask:long_name = "VIIRS Integer Cloud Mask" ;
                    Integer_Cloud_Mask:DataDescription = "VIIRS cloud mask bits 1 & 2 converted to
integer (0 = cloudy, 1= probably cloudy, 2 = probably clear, 3 = confident clear, -1 = no
result)" ;
                    Integer_Cloud_Mask:units = "none" ;
                    Integer_Cloud_Mask:_FillValue = -1b ;
                    Integer_Cloud_Mask:valid_min = 0s ;
                    Integer_Cloud_Mask:valid_max = 3s ;
            ubyte Quality_Assurance(number_of_lines, number_of_pixels, QA_dimension) ;
                    Quality_Assurance:long_name = "Quality Assurance for VIIRS Cloud Mask" ;
                    Quality_Assurance:units = "none" ;
                    Quality_Assurance:_FillValue = 0UB ;
                    Quality_Assurance:valid_min = 1 ;
                    Quality_Assurance:valid_max = 255 ;
    } // group geophysical_data
group: scan_line_attributes {
    variables:
            double scan_start_time(number_of_scans) ;
                    scan_start_time:long_name = "Scan start time (TAI93)" ;
                    scan_start_time:units = "seconds" ;
                    scan_start_time:_FillValue = -999.9 ;
                    scan_start_time:valid_min = 0. ;
                    scan_start_time:valid_max = 2000000000. ;
    } // group scan_line_attributes
}
```

Table 3. Bit locations for data in the MVCM 'Cloud_Mask' array

		Result
0	Cloud Mask Flag	0 = not determined 1 = determined
1–2	Unobstructed FOV Confidence Flag	00 = cloudy 01 = probably cloudy 10 = probably clear 11 = confident clear
Processing Path Flags		
3	Day / Night Flag	0 = Night / 1 = Day
4	Sun glint Flag	0 = Yes / 1 = No
5	Snow / Ice Background Flag	0 = Yes / 1 = No
6–7	Land / Water Background Flag	00 = Water 01 = Coastal 10 = Desert 11 = Land
	1-km Flags	
8	Spare	0 = Yes / 1 = No
9	Thin Cirrus Detected (solar)	0 = Yes / 1 = No
10	Snow cover from ancillary map	0 = Yes / 1 = No
11	Thin Cirrus Detected (infrared)	0 = Yes / 1 = No
12	Cloud Adjacency (cloudy, prob. cloudy, plus 1-pixel adjacent)	0 = Yes / 1 = No
13	Cloud Flag – Ocean IR Threshold Test	0 = Yes / 1 = No
14	Spare	0 = Yes / 1 = No
15	Spare	0 = Yes / 1 = No
16	High-Cloud Flag – 1.38 μm Test	0 = Yes / 1 = No
17	High-Cloud Flag – 3.9-12 μm Test (night only)	0 = Yes / 1 = No
18	Cloud Flag – IR Temperature Difference Tests	0 = Yes / 1 = No
19	Cloud Flag – 3.9-11 μm Test	0 = Yes / 1 = No
20	Cloud Flag – VNIR Reflectance Test	0 = Yes / 1 = No
21	Cloud Flag – VNIR Reflectance Ratio Test	0 = Yes / 1 = No
22	Clear-sky Restoral Test – NDVI in Coastal Areas	0 = Yes / 1 = No
23	Cloud Flag – Water 1.6 or 2.1 μm Test	0 = Yes / 1 = No
24	Cloud Flag – Water 8.6-11 μm Test	0 = Yes / 1 = No
25	Clear-sky Restoral Test – Spatial Consistency (ocean)	0 = Yes / 1 = No
26	Clear-sky Restoral Tests (polar night, land, sun glint)	0 = Yes / 1 = No
27	Cloud Flag – Surface Temperature Tests (water, night land)	0 = Yes / 1 = No
28	Spare	0 = Yes / 1 = No
29	Spare	0 = Yes / 1 = No

30	Cloud Flag – Night Ocean 11 µm Variability Test	0 = Yes / 1 = No
31	Cloud Flag – Night Ocean "Low-Emissivity" 3.9–11 µm Test	0 = Yes / 1 = No
I-Band/250-m Cloud Flag		
32	Element(1,1)	0 = Yes / 1 = No
33	Element(1,2)	0 = Yes / 1 = No
34	Element(1,3)	0 = Yes / 1 = No
35	Element(1,4)	0 = Yes / 1 = No
36	Element(2,1)	0 = Yes / 1 = No
37	Element(2,2)	0 = Yes / 1 = No
38	Element(2,3)	0 = Yes / 1 = No
39	Element(2,4)	0 = Yes / 1 = No
40	Element(3,1)	0 = Yes / 1 = No
41	Element(3,2)	0 = Yes / 1 = No
42	Element(3,3)	0 = Yes / 1 = No
43	Element(3,4)	0 = Yes / 1 = No
44	Element(4,1)	0 = Yes / 1 = No
45	Element(4,2)	0 = Yes / 1 = No
46	Element(4,3)	0 = Yes / 1 = No
47	Element(4,4)	0 = Yes / 1 = No

File Naming Conventions

Output file names follow the format listed below. An example MVCM VIIRS granule is shown.

CLDMSK_L2_VIIRS_SNPP.AYYYYDDD.HHMM.VVV.YYYYDDDHHMMSS.nc

The interpretation of this file name is as follows:

CLDMSK: Data product type

L2: Data product level (Level-2 pixel-level: **L2**; Level-3 global gridded: **L3**)

VIIRS: Sensor name (**MODIS**, **VIIRS**)

SNPP: Platform name (**Aqua**, **SNPP**)

AYYYYDDD: Data acquisition year (**YYYY**) and day of year (**DDD**)

HHMM: Data acquisition hour (**HH**) and minute (**MM**) start time, in UTC

VVV: Data version number

YYYYDDDHHMMSS: Data production date and time, in UTC

nc: Denotes NetCDF-4 file format

Validation

Cloud Detection

Since 2006, a very effective validation tool has been available for use with Aqua MODIS cloud algorithms. The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) (Winker, et al., 2007) on board the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) platform detects clouds with high accuracy (Vaughn, et al., 2009). From its launch in 2006 until mid-2018, the CALIPSO platform flew in formation as part of the "A-Train" constellation of satellites (Stephens et al., 2002), lagging Aqua by about 95 seconds. This resulted in both Aqua MODIS and CALIOP observing the same clouds and/or Earth surface nearly simultaneously. CALIOP observations are from nadir only with Earth locations that precess across collocated MODIS scans but do not include sun-glint regions or MODIS pixels from far-limb areas. Here, we compare the Aqua MODIS cloud mask (MYD35) and the MVCM Aqua MODIS to collocated CALIOP data (used as "truth"). This comparison poses a rather severe test of both the MYD35 and MVCM Aqua MODIS since the lidar is sensitive to very thin clouds (Holz et al., 2008). CALIOP data may also be collocated with MVCM VIIRS, though we cannot expect as many "high quality" collocated pixels available for analysis because the SNPP platform is not part of the A-train constellation. A five-minute filter is used on SNPP

VIIRS vs. CALIOP comparisons; if the time difference between the two observations is larger than five minutes, those observation pairs are not considered in the validation statistics.

Tables 4 and 5 show hit rates, false positive (false cloud), and false negative (missed cloud) rates for various scene types between CALIOP and Collection 6.1 Aqua MODIS ("MYD35" in the table), MVCM Aqua MODIS, and MVCM VIIRS for February and August of 2014. There is generally good agreement with CALIOP especially in non-polar regions. Also, note the increase in agreement from northern hemisphere winter to summer as the area of snow cover decreases greatly over the planet. Interestingly, the MVCM shows closer agreement to CALIOP than MYD35 for 60S–60N daytime land in February, and for global night, 60S–60N night, and 60S–60N night land in August. The former is due to improved snow detection vs. vegetation via spectral tests, while the latter is mainly due to reduction of false positives (labeling clear-sky pixels as cloudy).

Table 4. February 2014 overall hit rates, false positive (false cloud), and false negative (missed cloud) rates comparing Aqua MODIS cloud mask (MYD35), MVCM Aqua MODIS, and MVCM VIIRS to CALIOP cloud detection data

Feb 2014 MYD35, MVCM MODIS, and MVCM VIIRS vs. CALIOP Cloud Detection						
Scene Type	Hit Rates (%)			False Pos. / False Neg. (%)		
	MYD35	Aqua MODIS	VIIRS	MYD35	Aqua MODIS	SNPP VIIRS
Global	88.0	86.5	86.3	12.9/11.5	16.1/12.2	14.1/13.6
60S–60N	90.1	89.6	89.6	15.0/7.4	16.7/7.4	15.7/7.8
Global Day	90.2	89.6	88.9	11.6/8.8	15.1/7.6	14.0/9.5
Global Night	86.1	83.8	84.0	14.4/13.8	17.2/15.8	14.2/16.9
60S–60N Day	91.1	90.5	90.5	13.1/6.7	18.4/4.6	16.4/5.6
60S–60N Night	89.2	88.7	88.8	17.4/8.2	14.6/9.9	14.7/9.8
60S–60N Water Day	92.6	91.1	91.1	11.1/5.7	23.1/2.2	20.1/3.8
60S–60N Water Night	90.2	90.2	89.8	17.8/7.2	15.7/8.0	15.5/8.5
60S–60N Land Day	87.0	88.9	88.9	17.0/9.7	9.5/12.4	10.6/11.5
60S–60N Land Night	86.2	84.7	85.6	16.6/11.6	12.7/17.2	13.1/15.2
60S–60N Desert Day	85.8	87.5	88.3	14.4/13.8	10.2/16.4	10.1/14.5
60S–60N Desert Night	83.6	82.7	82.9	15.4/18.0	11.6/25.2	10.2/29.4
Polar Day	88.1	87.4	85.1	8.3/14.4	7.9/15.7	8.5/19.5
Polar Night	80.7	75.3	76.0	9.6/23.6	21.5/26.2	13.6/29.6

Table 5. August 2014 overall hit rates, false positive (false cloud), and false negative (missed cloud) rates comparing Aqua MODIS cloud mask (MYD35), MVCM Aqua MODIS, and MVCM VIIRS to CALIOP cloud detection data

Aug 2014 MYD35, MVCM MODIS, and MVCM VIIRS vs. CALIOP Cloud Detection						
Scene Type	Hit Rates (%)			False Pos. / False Neg. (%)		
	MYD35	Aqua MODIS	VIIRS	MYD35	Aqua MODIS	VIIRS
Global	88.3	88.1	87.5	12.8/11.2	15.1/10.4	13.5/12.1
60S–60N	91.2	91.1	90.4	11.4/7.4	13.4/6.6	13.3/7.7
Global Day	91.6	91.0	89.0	10.4/7.4	15.6/5.8	15.8/8.8
Global Night	85.3	85.5	86.2	14.9/14.6	14.5/14.5	11.3/15.0
60S–60N Day	91.8	91.3	90.1	9.4/7.6	15.2/5.1	15.8/6.6
60S–60N Night	90.6	90.9	90.7	13.6/7.3	11.4/8.0	10.5/8.7
60S–60N Water Day	92.2	91.5	90.2	12.3/5.7	21.1/2.7	21.6/4.4
60S–60N Water Night	90.4	90.4	90.2	17.3/6.9	15.1/7.7	13.2/8.7
60S–60N Land Day	90.5	90.7	89.8	4.4/14.4	4.7/13.7	6.4/13.9
60S–60N Land Night	91.3	92.3	92.1	8.4/9.0	6.0/9.5	6.8/9.0
60S–60N Desert Day	91.5	91.6	90.8	1.1/28.6	1.5/27.1	2.7/26.6
60S–60N Desert Night	91.4	91.8	93.0	3.6/20.7	3.0/20.6	2.1/20.1
Polar Day	91.3	90.4	86.2	14.2/7.1	17.3/7.4	15.5/13.3
Polar Night	76.3	75.4	77.9	17.1/26.7	21.3/25.8	13.1/25.7

MODIS and VIIRS Continuity

A time series of daily cloud fractions taken from pixel-level MVCM Aqua MODIS and MVCM SNPP VIIRS results was created over the North Pacific region (0–50N, 160E–128W). Figures 2 and 3 show time-series plots of Aqua MODIS daytime cloud fraction, and Aqua MODIS minus SNPP VIIRS daytime cloud fraction differences, respectively, over the period from March 2012 through November 2018. Cloud fractions generally vary between 60 and 80 percent, and are higher during the winter months and lower during the summer, especially late in the season. The MODIS data yields slightly higher cloudiness overall with a bias of 1.92 percent relative to VIIRS. Surprisingly, the differences also show seasonality where the largest positive differences are during winter, and lowest positive or slightly negative differences are seen during the summer season. Despite the seasonality, the time series of differences is very stable across the time domain, and in the aggregate have a Gaussian distribution with mean minus median values of only 0.03 percent. We speculate that the reason for the positive bias between MODIS and VIIRS is the larger

footprint of MODIS. One possible reason for the seasonal aspect of the differences is the relative abundance of open-cell marine low clouds during the winter compared to higher occurrence of closed-cell stratus during summer. With a smaller FOV, the VIIRS is perhaps more likely to detect clear skies within open-cell clouds than is MODIS.

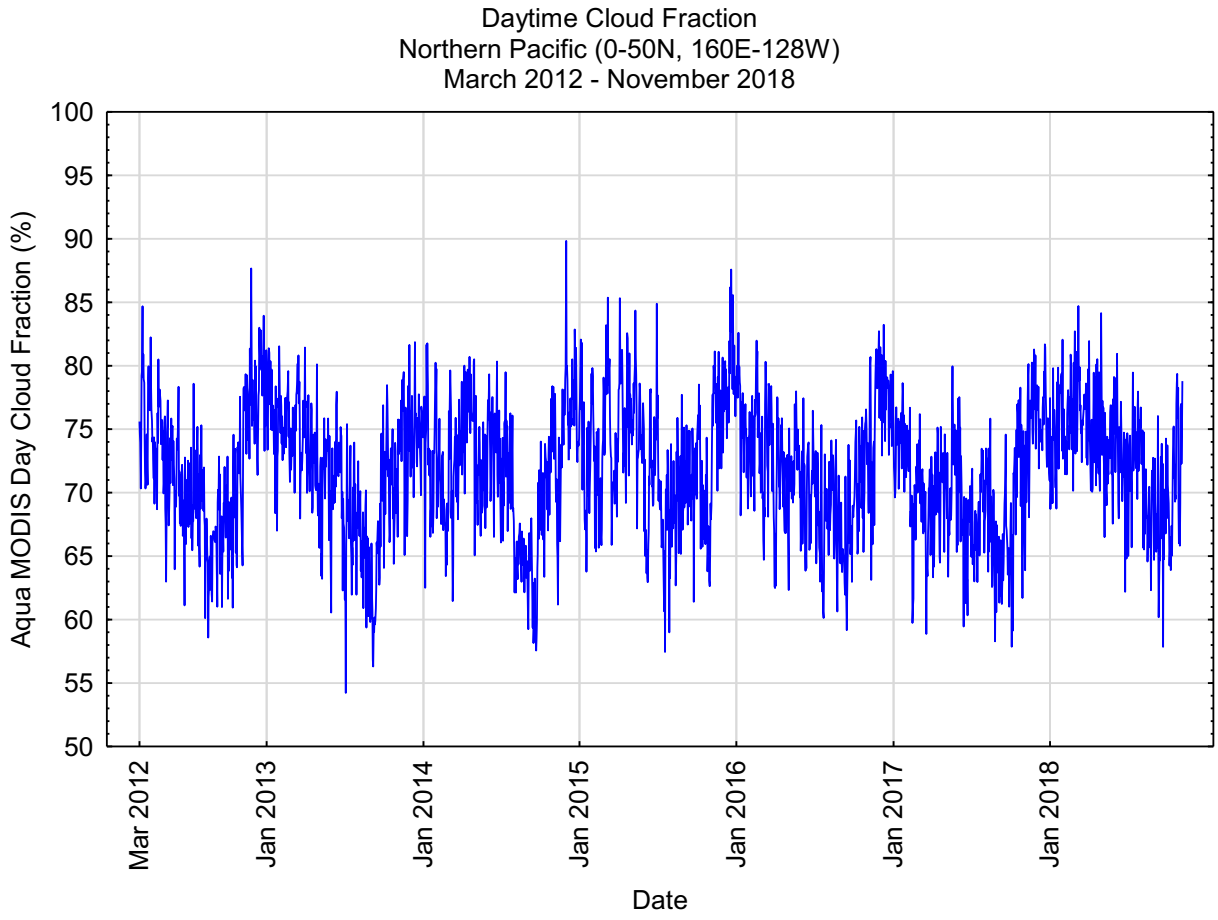

Figure 2. Time series of daily mean cloud fractions taken from MVCM Aqua MODIS pixel-level results

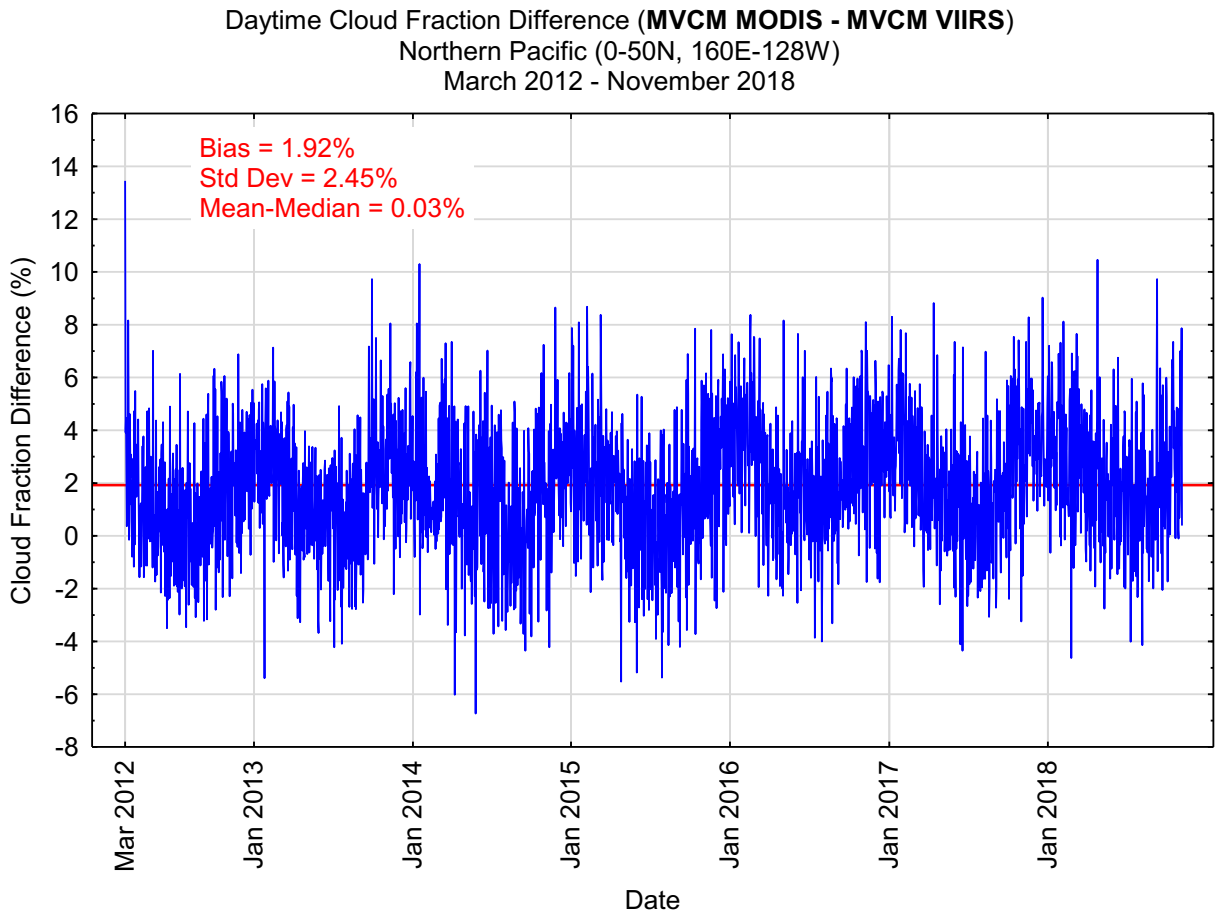

Figure 3. Time series of daily mean cloud fraction differences between MVCM Aqua MODIS and MVCM SNPP VIIRS

Use of Alternate Clear vs. Cloudy Thresholds

The MVCM output will include the clear-sky confidence, or Q value, for each pixel (see Contents section above). This section details the relationships between Q and the success and failure rate in clear and cloudy-sky detection, and how users may choose alternate Q values for clear vs. cloudy discrimination. These relationships are shown in terms of Receiver Operator Characteristics (ROC) graphs (Fawcett, 2005). In plots of satellite imager cloud detection ROC curves, the y-axis typically represents the "true positive rate" and the x-axis shows the "false positive rate". The first case is when "truth" (in our case, from space-borne CALIOP lidar) indicates cloud, and the imager algorithm also indicates cloud. The latter is when truth indicates clear skies but the imager cloud detection algorithm denotes clouds. Both axes have values from 0.0 to 1.0 that define the frequency range of both conditions. A perfect algorithm would show one point on the plot; a true positive rate of 1.0 and a false positive rate of 0.0. Typically, however, there are multiple points on these plots that correspond to differing modes of algorithm operation. Connecting the

points results in a curve that defines the ROC. The purpose of the ROC curve is to ascertain the optimal operational mode, maximizing the true positive rate while minimizing the false positive rate.

Figure 4 displays an example below. Here the points on the curve define "Q", the calculated confidence of clear sky from the MVCM cloud detection algorithm. Values of Q increase from left to right (0 to 1). The ROC curve shows how cloud detection (true positive rate) vs. false clouds (false positive rate) would change in the aggregate as various values of Q are used as thresholds to discriminate between clear and cloudy pixels. In the MVCM method, all values of clear-sky confidence equal to or below the threshold Q value are labeled *cloudy*, and all values above that threshold are labeled *clear*.

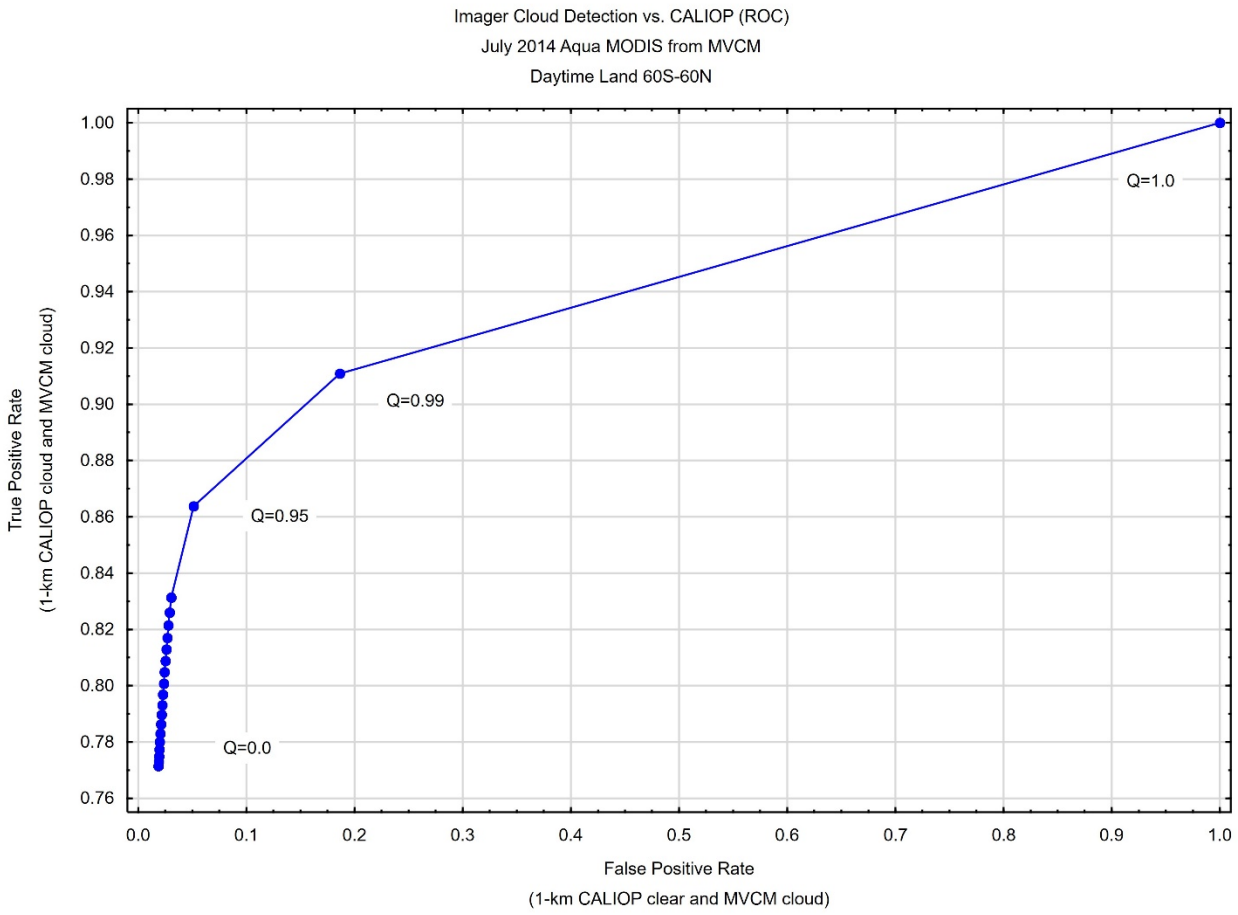

Figure 4. ROC curve for daytime land scenes from MVCM Aqua vs. CALIOP lidar in July, 2014

Daytime Land Example

From the curve in Figure 4, we see that a Q threshold of 0 (all pixels with clear-sky confidences of 0.0 are cloudy, and all other pixels are clear) means that about 76.5 percent of all cloudy pixels will be correctly labeled, and only about 2 percent of clear pixels will be mislabeled. Moving through the remaining points on the ROC curve, we find the proportion of correctly identified clouds increases as the Q threshold

increases because many clouds are associated with non-zero clear-sky confidences. The false-cloud rate also increases but not as quickly near the left side of the graph because more clear pixels have clear-sky confidence values closer to 1.0 than to 0.0. As one approaches the right-hand side of the curve, proportions of both correct *and* incorrect cloudy designations increase, with the incorrect ones increasing more rapidly to the right of Q=0.95, the MVCM (and MODIS cloud mask) clear vs. cloudy threshold.

Why do both rates increase with increasing Q thresholds? Since Q (cloud determinations) ≤ Q (threshold) < Q (clear determinations), as Q (threshold) nears 1.0, almost all truth clouds will be *detected*, including even some that are too optically thin or physically small for actual imager detection. At the same time, more and more truly clear pixels will be incorrectly labeled cloudy. And, since the pixels with calculated Q values near 1.0 are more likely to be clear than cloudy, the false positive rate increases more rapidly than the true positive rate. Measurement and algorithm parameter uncertainties also play a larger role as calculated Q values and the Q threshold approach 1.0. Clear pixels, pixels containing very thin clouds, and pixels in or near cloud edges can all have Q values at or near 1.0.

Daytime Ocean Example

Figure 5 shows a similar plot but for daytime ocean instead of land. Due to the consistently dark ocean background, clouds are more easily detected over water than land. For a Q threshold of 0.0, already almost 93 percent of all clouds are correctly labeled; however, the proportion of false clouds is about 11 percent, higher than the land case. This ROC curve suggests that a lower Q threshold could be used to separate clear from cloudy pixels in the MVCM. On the other hand, due to its high sensitivity to oceanic clouds, the MVCM is capable of identifying small-sized cumulus and thin cirrus that is simply not possible over land, and we can take advantage of this capability. The dark background can also lead to aerosols being mislabeled as clouds, and our truth data does not distinguish between clear skies and clear skies that contain aerosols. Users with a somewhat higher tolerance for aerosol or other “minor” cloud contamination could perhaps find a lower Q value to use as a clear vs. cloudy threshold, or indeed, all four cloud detection categories output by the cloud mask. The algorithm currently uses Q > 0.99 (confident clear), Q > 0.95 (probably clear), Q > 0.66 (probably cloudy), and Q ≤0.66 (cloudy).

Another way to look at cloud detection efficiency is seen in Figure 6, where the plot components have been rearranged so that the Q thresholds are on the x-axis, and the true-positive and false-positive rates are on the left and right-hand axes, respectively. The discontinuity in both rates between Q threshold 0.65 and 0.70 is due to clear-sky restoral tests. Q values are arbitrarily changed from < 0.67 to ≥ 0.67 if there are conflicting clear and cloudy-sky signals in the radiance data. The left-hand side of both plots is almost flat because a large majority of cloudy ocean pixels have Q=0.0, and a great many clear pixels have

Q=1.0. Clear-sky restoral tests also play a role by removing some Q values < 0.67 and placing them farther to the right on the x-axis.

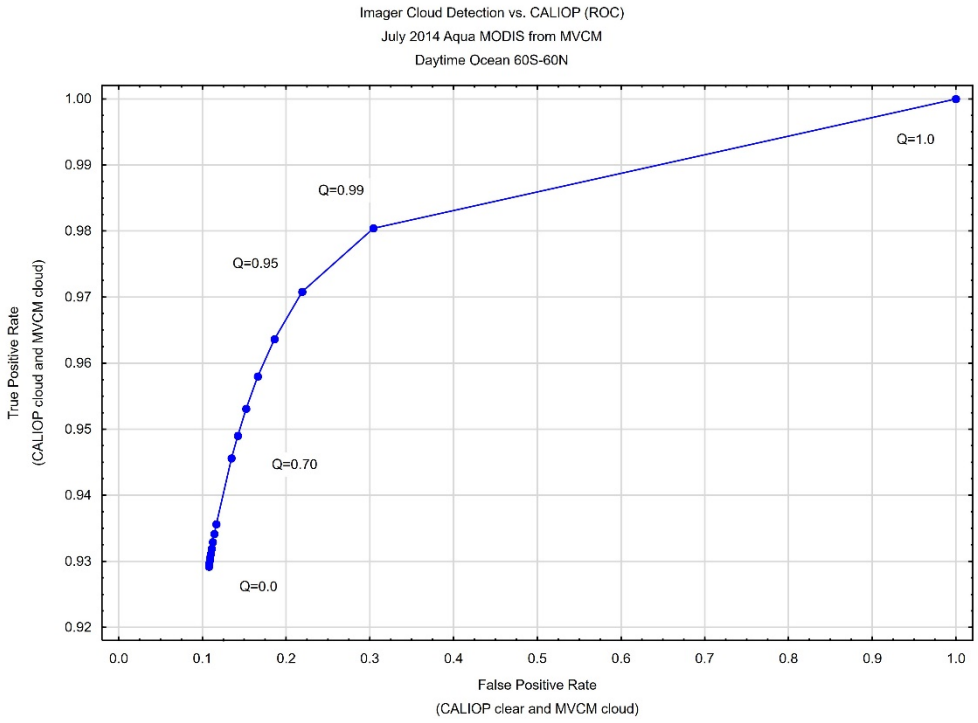

Figure 5. ROC curve for daytime ocean scenes from MVCM Aqua vs. CALIOP Lidar in July, 2014

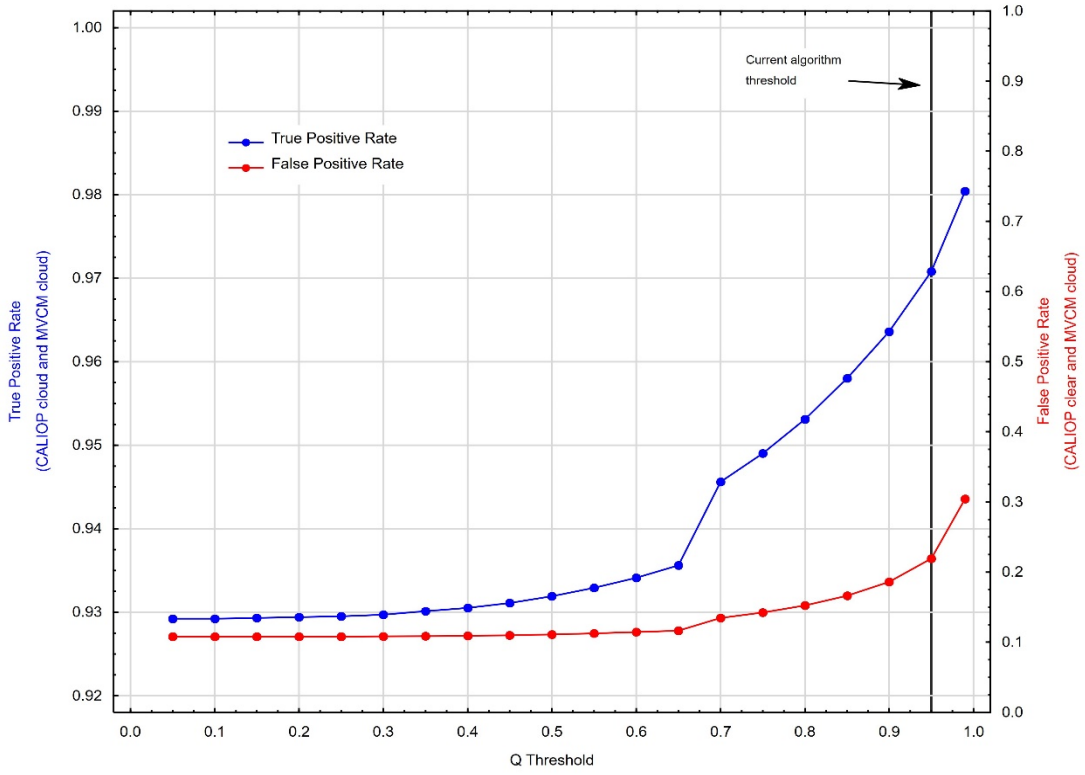

Figure 6. True and false positive rates as a function of Q thresholds. MVCM vs. CALIOP Lidar in July 2014

Acknowledgements

The research was supported by NASA Contract NNX15AG12G. Algorithm development support was performed at the Atmosphere SIPS under NASA contract NNG15HZ38C.

References

Ackerman, S.A., K. I. Strabala, W. P. Menzel, R. A. Frey, C. C. Moeller, and L. E. Gumley, Discriminating clear sky from clouds with MODIS, *J. Geophys. Res.*, *103*(D24), 32141-32157, 1998.

Ackerman, S. A., R. A. Frey, K. I. Strabala, Y. Liu, L. E. Gumley, B. A. Baum, and W. P. Menzel, Discriminating clear-sky from clouds with MODIS Algorithm Theoretical Basis Document (MOD35), Ver. 6.1, October 2010, MODIS Atmosphere Web Site, 117 pp., 2010.

Chen, S. and T. Zhang, An improved cloud masking algorithm for MODIS ocean colour data processing, *Remote Sens. Letters* 6.3, pp. 218-227, 2015.

Fawcett, T, An introduction to ROC analysis, *Pattern Recognition Letters*, 27, 861-874, 2005.

Holz, R. E., S. A. Ackerman, F. W. Nagle, R. A. Frey, S. Dutcher, R. E. Kuehn, M. Vaughan, and B. A. Baum, Global MODIS cloud detection and height evaluation using CALIOP. *J. Geophys. Res*., 113, D00A19, doi:10.1029/2008JD009837, 2008.

Klein, A. G., D. K. Hall, and G. A. Riggs, Improving snow cover mapping in forests through the use of a canopy reflectance model, *Hydrological Processes* 12.10, pp. 1723-1744, 1998.

Moody, E. G., M. D. King, and S. Platnick, Spatially complete global spectral surface albedos: value-added datasets derived from Terra MODIS land products. IEEE Trans. *Geosci. Rem. Sci.* 43, 144-158, 2005.

Stephens, G.L., D.G. Vane, R.J. Boain, G.G. Mace, K. Sassen, Z. Wang, A.J. Illingworth, E.J. O'Connor, W.B. Rossow, S.L. Durden, S.D. Miller, R.T. Austin, A. Benedetti, C. Mitrescu, THE CLOUDSAT MISSION AND THE A-TRAIN. *Bull. Amer. Meteor. Soc.,* 83, 1771–1790, https://doi.org/10.1175/BAMS-83-12-1771, 2002.

Vaughan, M., K. A. Powell, D. M. Winker, C. A. Hostetler, R. E. Kuehn, W. H. Hunt, B. J. Getzewich, S. A. Young, Z. Liu, and M. J. McGill, Fully automated detection of cloud and aerosol layers in the CALIPSO Lidar measurements. *Journal of Atmospheric and Oceanic Technology*, Volume 26, 2034–2050, 2009.

Weinreb, M. P., R. Xie, J. H. Lienesch, and D. S. Crosby, Destriping GOES images by matching empirical distribution functions. *Remote Sens. Environ*., 29, 185-195, 1989.

Winker, D. M., W. H. Hunt, and M. J. McGill, Initial performance assessment of CALIOP, *Geophysical Research Letters*, Volume 34, L19803, doi:10.1029/2007GL030135, 2007.

Xiong, X., J. Butler, K. Chiang, B. Efremova, J. Fulbright, N. Lei, J. McIntire, H. Oudrari, J. Sun, Z. Wang, and A. Wu, VIIRS on-orbit calibration methodology and performance, *J. Geophys. Res. Atmos.,* 119, pp. 5065–5078, doi:10.1002/2013JD020423, 2014.